SOFTWARE ARTS TECHNICAL NOTES, BY THE CREATORS OF VISICALC

A Publication of
Software Arts, Inc.

SATN

THE JOURNAL FOR VISICALC USERS

Vol. I, No. 5
May/June 1982

IN THIS ISSUE:

About This Issue 1
Three Forecasting Models 1
Interesting Comparisons 6
Tee Shirts, Drive Shafts, and Bonds 9

. . . Plus VisiCalc Pointers

ABOUT THIS ISSUE

Fall brings elections, and in this issue Paul Hoffman takes a look at some of the ways that the VisiCalc® program can be used in analyzing voter preferences. If analyzing elections is not one of the ways you spend your time with your personal computer, the techniques and methods shown in *Three Forecasting Models* can be used for market research or commodities trading.

In *Interesting Comparisons* we look at the problem of comparing interest rates that are compounded at different frequencies. Although the thrust of the article is primarily financial, the technique used has many other applications. One of them, population growth studies, is mentioned in the article itself, but in addition to that, the general formulas we use can be applied to problems of comparing rates of growth and decline in business, and in the social, biological, and physical sciences.

In an earlier issue we alluded to using a mathematical tool called "modulo" for the creation of a part number. In *Tee Shirts, Drive Shafts, and Bonds,* we have picked up that thread. The technique described in the article is useful not only for creating part numbers for bills of material, but can be adapted for any application in which you need to combine a large number of variables into a single number that uniquely describes each possible combination. Cross correlating the variables in a market survey or opinion poll is one of the problems that comes to mind as an alternative use of the technique.

The next issue of SATN will mark the end of our first year of publication and the addition of two new features to look forward to. Beginning with the July/August issue we will include selected overviews of hardware and software designed to complement the VisiCalc program. Those of you who are regular readers of general personal computer magazines may already know or own some of these products. For those who do not read those journals, we hope to provide a source of information on the availability of products that may be useful in conjunction with the VisiCalc program.

We are not the only ones who write about how to apply the VisiCalc program, so in the September/October issue we will begin a guide to articles about the program that have appeared elsewhere. We have no ambitions to compile a complete biography, but we will reference those articles and books that contain useful applications and techniques.

Assuming that there is enough activity on those fronts, we will make them regular, alternating features. If not, they will appear irregularly.

THREE FORECASTING MODELS

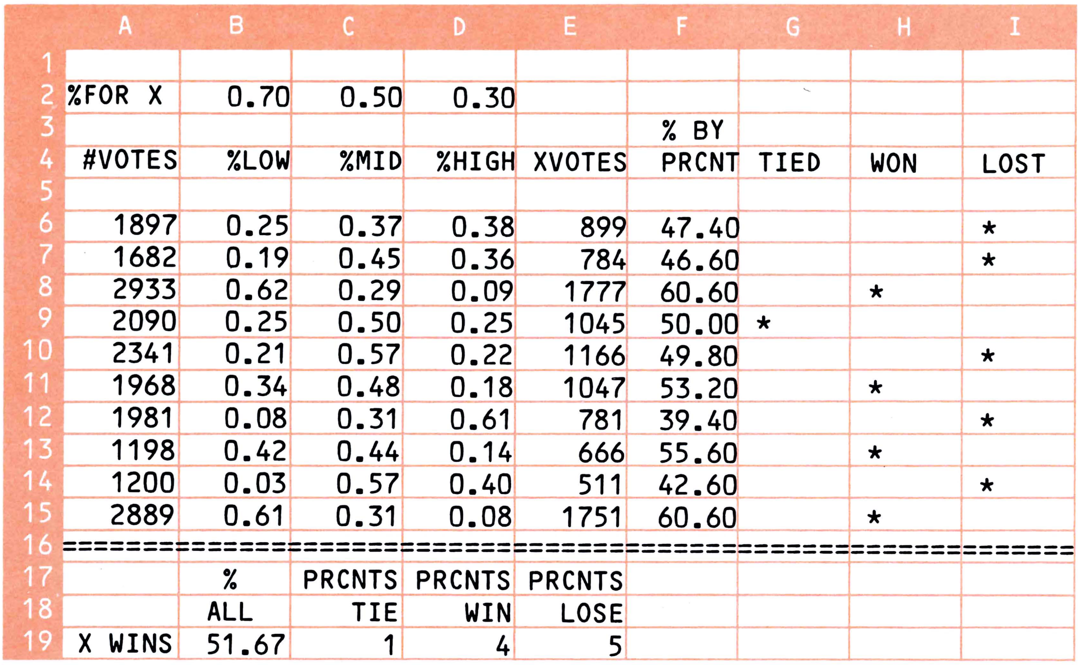
By Paul Hoffman

Forecasting, it has been said, can be a tricky business, especially if it's about the future. Predicting the outcome of an election is one of the trickier sorts of forecasting. Using the VisiCalc® program can't make the process any more accurate, but it can help the forecaster see how results change as conditions or assumptions about an election are altered.

We'll look at three specific models: a group weight, a winner-takes-all, and a trend analysis. Each is used in political forecasting, but each can also be applied easily to marketing strategy or commodity futures planning. Since the VisiCalc program allows us to make calculations on large numbers of variables, these models can be extended and combined to give greater depth to the forecasting process.

A Group Weight Model

Group weight models are based on our ability to identify groups of persons and one or more factors that influence their behavior. By assigning numerical values to those factors, we weight the group and pre-



	A	B	C	D	E	F	G	H	I
1									
2	%FOR X	0.70	0.50	0.30					
3						% BY			
4	#VOTES	%LOW	%MID	%HIGH	XVOTES	PRCNT	TIED	WON	LOST
5									
6	1897	0.25	0.37	0.38	899	47.40			*
7	1682	0.19	0.45	0.36	784	46.60			*
8	2933	0.62	0.29	0.09	1777	60.60		*	
9	2090	0.25	0.50	0.25	1045	50.00	*		
10	2341	0.21	0.57	0.22	1166	49.80			*
11	1968	0.34	0.48	0.18	1047	53.20		*	
12	1981	0.08	0.31	0.61	781	39.40			*
13	1198	0.42	0.44	0.14	666	55.60		*	
14	1200	0.03	0.57	0.40	511	42.60			*
15	2889	0.61	0.31	0.08	1751	60.60		*	
16	==========	==========	==========	==========	==========	==========	==========	==========	==========
17		%	PRCNTS	PRCNTS	PRCNTS				
18		ALL	TIE	WIN	LOSE				
19	X WINS	51.67	1	4	5				

FIGURE 1. *The TIED, WON, and LOST columns can be calculated with @IF statements that differ only in the comparison operator used. Enter:*

```
>G6:@IF(E6/A6=.5,1,0)
>H6:@IF(E6/A6>.5,1,0)
>I6:@IF(E6/A6<.5,1,0)
```

dict changes in the outcome of a campaign. And by varying the numerical values assigned, we can determine the influence of the factors on the overall results. When the groups and the factors affecting them are accurately identified, a group weight model can help the political candidate or marketing manager identify and concentrate on those activities that will produce the largest gain.

Group weight models are also useful for developing marketing strategies and identifying areas where a small reallocation of resources may produce a large sales increase.

In our example, votes will be grouped by precinct. We will use household income as the factor to indicate voting preference for candidate X. The factor identified in this example could as easily have been age or occupation.

SATN is published bi-monthly by Software Arts, Inc., P.O. Box 494, Cambridge, Massachusetts 02139. Address all editorial correspondence to the Editorial Department, SATN, P.O. Box 494, Cambridge, MA 02139. Requests for special permission or bulk orders should be addressed to the publisher.

Domestic subscriptions: 6 issues, $30.00; outside U.S.A., $50.00. Single copy price, $6.00. Send subscriptions, change of address, and fulfillment questions to SATN Subscriptions, P.O. Box 815, Quincy, MA 02169.

Copyright © 1982, Software Arts, Inc. All rights reserved. This periodical may not be copied or reproduced in any fashion either in whole or in part without the prior written consent of Software Arts, Inc.

Software Arts, Inc. makes no warranty, express or implied, with respect to the quality, accuracy or freedom from error of the information contained in SATN, the journal for VisiCalc® users, and specifically disclaims all liability for any damages resulting from the use of such information.

Apple is a registered trademark of Apple Computer, Inc. Atari is a trademark of Atari Inc. Commodore is a trademark of Commodore International. DIF is a trademark of Software Arts, Inc. HP-83/85 and HP-125 are trademarks of Hewlett-Packard Company. IBM is a registered trademark of International Business Machines Corporation. TRS-80 is a trademark of Radio Shack Division of Tandy Corp.

SATN is a trademark of Software Arts, Inc.
VisiCalc is a registered trademark of VisiCorp/Personal Software Inc.

Set up your sheet with the commands:

```
/GOR
/GC7
/GF$
```

to calculate across rows, set the column width to seven characters, and display two decimal places.

Figure 1 shows how a group weight model could be entered on a VisiCalc sheet. In row 2, %FOR X, we have entered predictions of the percent of voters in each category who will vote for candidate X. We have predicted that X will get 70% (0.70) of the low income vote, 50% (0.50) of the middle income vote, and only 30% (0.30) of the high income vote.

The #VOTES column contains the voting populations by precincts. The columns LOW, MID, and HIGH contain data of the proportions of low, middle, and high income households in each precinct. These columns contain the "givens" of an election, with HIGH containing the only calculated values. They are calculated simply by subtracting the values entered in columns B and C from 1. Enter the values shown in columns A . . . C, rows 6 . . . 15. Because we will later edit this model and make use of column A in another model, the values in that column should be formatted with /FI to show integers. That can be done by replicating a local format command into the empty cells. Enter:

```
>A6:/FI/R:A7...A15
```

To calculate the HIGH column, enter:

```
>D6:1-B6-C6®/R:D7...D15:RR
```

To find X VOTES, which is the group weight total for a precinct, we calculate the number of votes X will gather from each income group in each precinct and total them into column E, which we have formatted with /FI. The general formula for that series of calculations is:

#VOTES * ((%LOW * proportion low favoring X) +
(%MID * proportion mid favoring X) +
(%HIGH * proportion high favoring X))

As entered on the sheet, and replicated down the column it is:

```
>E6:/FI+A6*(B6*B2+(C6*C2)+(D6*D2))
>E6:/R:E7 . . . E15:RRNRNRN
```

In column F, % BY PRCNT, are the percentages of voters favoring X in each precinct. The calculation of that column is:

```
>F6:+E6/A6*100®/R:F7 . . . F15:RR
```

Columns G . . . I contain expressions using @INT formatted with /F∗. As their labels indicate, they show the precincts in which X is tied, winning, or losing. (If your version of the VisiCalc program has @IF, see the caption to Figure 1 for the appropriate expressions.) Enter:

```
>G6:/F*@INT(E6/A6*@INT(A6/E6))
>H6:/F*@INT(E6/A6+.5)-G6
>I6:/F*@INT(1/(E6/A6+.5))-G6
>G6:/R . . . I6:G7 . . . G15:RRRRRRRRRR
```

As long as E6 or A6 does not equal zero, those entries will evaluate to one or zero and display stars identifying the precincts in which XVOTES are respectively equal to, greater than, or less than 50%.

The total of the votes received by X in all precincts is shown as a percentage of all votes cast in the X WINS row at B19 under % ALL. Enter:

```
>B19:@SUM(E6 . . . E16)/
  @SUM(A6 . . . A16)*100
```

The model shows that X will receive 51.67% of the votes.

Ties, wins, and losses are included in the X WINS row with the entries:

```
>C19:@SUM(G6 . . . G16)
>C19:/R:D19 . . . E19:RR
```

By changing the weight assigned to any group we can find the one that, given the composition of the various precincts, is most influential. Assume, for example, that by changing position on taxation, X can gain votes in one of the income groups without losing ground among voters in the other groups. Let's also say the largest change that can be made is 5%.

To compare the alternatives, go to B2 and change .7 to .75. That results in a total vote of 53.35% for X, eliminates the tie, and gives X six wins and four losses. Then change B2 back to .7, and change D17 to .35. The total vote for X is now 52.93%, indicating that the first position will win slightly more support. If the 5% change is made in the weight of the middle income voter group, however, X's vote total rises to 53.74%, an indication that on the issue of taxation, X would be wise to concentrate on the middle income group to gain ground on the opposition.

Once you have worked out this model, save the sheet, as we will now blank all but column A. After saving your sheet, enter:

```
>B1:/R . . . I1:B2 . . . B20
```

Now save column A, giving it the filename PRCNTS. Enter:

```
>A1:/SSPRCNTS
```

We will return to PRCNTS after developing a trend analysis model. First, however, we shall look at a simpler model.

A Winner-Takes-All Model

The American political process is such that some candidates are not elected by a simple majority or a plurality of the voters, but by a majority or plurality of electors in a winner-takes-all model.

The Electoral College system, under which the President and the Vice President of the United States are elected, is one example of a winner-takes-all model.

KEY TO CONVENTIONS USED IN SATN

The VisiCalc® program is available on personal computers manufactured by Apple, Atari, Commodore, Hewlett-Packard, IBM, and Radio Shack. To reduce any confusion arising from the differences among the keyboards of those computer systems, SATN uses a standard symbol and name to represent certain commonly used keys.

The symbol ® and the word RETURN are used as synonyms for the keys known as ENTER, ENDLINE, and RETURN. The symbol ^ is used to represent an exponentiation operator.

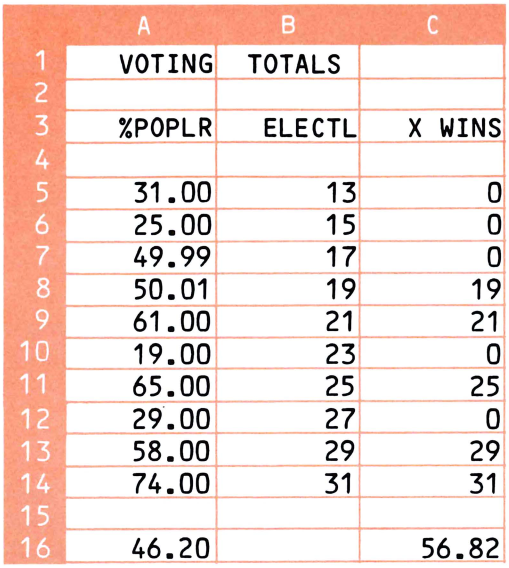
To make VisiCalc documentation more readable, the program was designed so that the colon (:) has the same effect as RETURN in some commands, as when ending the > command or separating various elements of the replicate command.

In examples of what to type, type exactly what you see and press RETURN at the end of every line. Users of the HP-83/85 may find it necessary to add parentheses to some of the expressions included in the text. For example, 3+4∗5 should be (3+4)∗5 on an HP-83/85.

Throughout SATN, one typeface represents both VisiCalc input and output (what you should type and what you will see on the screen). The typeface looks like this:

```
This line is typeset in OCR-B.
So is this: +A1+B1*3.25
```

To allow you to type the formulas directly on a VisiCalc sheet, examples in SATN specify coordinates, like A1, rather than the names of the values they represent. When actually typing a formula, you can use the arrows to point rather than using the > command.



	A	B	C
1	VOTING	TOTALS	
2			
3	%POPLR	ELECTL	X WINS
4			
5	31.00	13	0
6	25.00	15	0
7	49.99	17	0
8	50.01	19	19
9	61.00	21	21
10	19.00	23	0
11	65.00	25	25
12	29.00	27	0
13	58.00	29	29
14	74.00	31	31
15			
16	46.20		56.82

FIGURE 2. *The @IF expression to calculate X WINS is* >C5:@IF(A5/100<.5,0,B5), *which should be replicated relatively into C6 . . . C14.*

By winning a majority or plurality of the votes in a state, one candidate wins, theoretically, all of the electoral votes for that state. Thus, we have to determine whether candidate X's popular vote is less than 50%, and so gives him none of a state's electoral votes, or more than 50%, and all the state's electorial votes. To do so, we use the @INT function of VisiCalc.

Starting with a clear sheet, in column A, under %POPLR we enter percentage of the popular votes received by candidate X. Column B. ELECTL contains the number of electors received for winning a particular state. Enter the labels and values in columns A and B and in C3 as shown in Figure 2.

The first part of the calculation of X WINS, the number of electoral votes X has won, consists of dividing the percentage values in column A by 100, adding .5, and taking the integer portion of the result, which will always be 0 or 1. We then multiply the number of electoral votes by that result, to determine the number of electors that X has garnered. The first calculation is entered on the sheet at C5, and then replicated down the column as follows:

```
>C5:@INT(A5/100+.5)*B5
>C5:/R:C6 . . . C14:RR
```

On some versions of the VisiCalc program you can use the @IF function and a logical value to get the same result. The formula for those versions would be:

```
>C5:@IF(A5/100<.5,0,B5)
```

or:

```
>C5:@IF(A5/100>=.5,B5,0)
```

either of which is then replicated down the sheet as above.

To compare the difference between a candidate's popular vote and his or her electoral vote, we can average columns A and C to show the difference between the two. Enter:

```
>A16:@AVERAGE(A5 . . . A14)
```

and:

```
>C16:@SUM(C5 . . . C14)/
   @SUM(B5 . . . B14)*100
```

Although candidate X only got 46.2% of the popular vote, he or she received 56.82% of the electoral vote.

A Trend Analysis Model

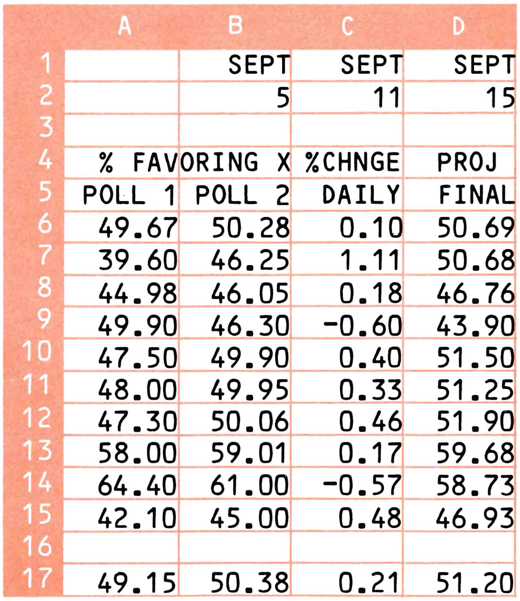
So far, the models we've shown have used static data. A trend analysis model analyzes changing data and can show us what should happen if a trend is followed to the day of an election.

Short-term trends can often be predicted using a linear model based on two or more points. As an example, we will use two popularity polls showing precinct-by-precinct changes in X's standing with the voters, and predict the outcome of the election based on X's standing at the time of the second poll.

Set up a clear sheet with the commands:

```
/GC7
/GOR
/GF$
```

to set the column width to seven characters, calculate across rows, and display to two decimal places.



	A	B	C	D
1		SEPT	SEPT	SEPT
2		5	11	15
3				
4	% FAVORING X		%CHNGE	PROJ
5	POLL 1	POLL 2	DAILY	FINAL
6	49.67	50.28	0.10	50.69
7	39.60	46.25	1.11	50.68
8	44.98	46.05	0.18	46.76
9	49.90	46.30	-0.60	43.90
10	47.50	49.90	0.40	51.50
11	48.00	49.95	0.33	51.25
12	47.30	50.06	0.46	51.90
13	58.00	59.01	0.17	59.68
14	64.40	61.00	-0.57	58.73
15	42.10	45.00	0.48	46.93
16				
17	49.15	50.38	0.21	51.20

FIGURE 3. *The advantage of a %CHNGE DAILY column is that it allows us to identify quickly those places where X is losing ground. If that step is not necessary, the PROJ FINAL column can be moved to the left and calculated with the expression,* >C6: + B6 – (B6 – A6/(B2 – C2) * (D2 – C2)). *That expression is replicated into C7 . . . C15 with the value references set to RRRNNNN.*

In our example, election day is September 15, and the popularity polls are taken six days apart, on September 5th and 11th. On the sheet shown in Figure 3, we have entered the results of the two polls in columns A and B. In column C the daily rate at which the polls changed is calculated for each precinct. In general terms the expression is:

$$\frac{\textit{Results of 1st Poll — Results of 2nd Poll}}{\textit{Days between Polls}}$$

On the sheet we have put the dates of the polls in row 2 as values, thereby allowing us to use the dates in our calculations by referring to them in the expressions we enter. (Note that those cells have been locally formatted with /FI.) Then, if the first poll were taken on September 3rd rather than the 5th, the formula entered at C6 will automatically take into account the number of days between polls. Our general formula to calculate the daily rate of change is entered on the sheet, and replicated down the column as:

```
>C6:+A6-B6/(B2-C2)®/R:C7 . . . C15:RRNN
```

We assume that the daily rate of change between September 5th and September 11th will remain constant until the election, meaning that all that remains is to multiply the rate by the number of days from the first poll to election day and add the result to X's standings as of the date of the first poll.

Enter:

```
>D6:+C6*(D2-B2)+A6
>D6:/R:D7 . . . D15:RNNR
```

On row 17 we calculate X's overall averages with @AVERAGE. Enter and replicate:

```
>A17:@AVERAGE(A6 . . . A15)®
>A17:/R:B17 . . . D17:RR
```

The percentages derived on row 17, we should point out, do not mean that X will poll 51.20% of the votes cast on election day. This model only shows a comparison of percentages, not of votes. If all precincts were the same size, then X's 51.20% projected for election day might translate into an equal percentage of all votes cast. Since the precincts are of different sizes, it won't. But we can calculate what these changes in X's popularity could mean by transferring the numbers of voters in each precinct from the sheet where we calculated the group weight model.

With the cursor at A1, insert a column. Enter:

```
>A1:/IC
```

Now load the file PRCNTS. Enter:

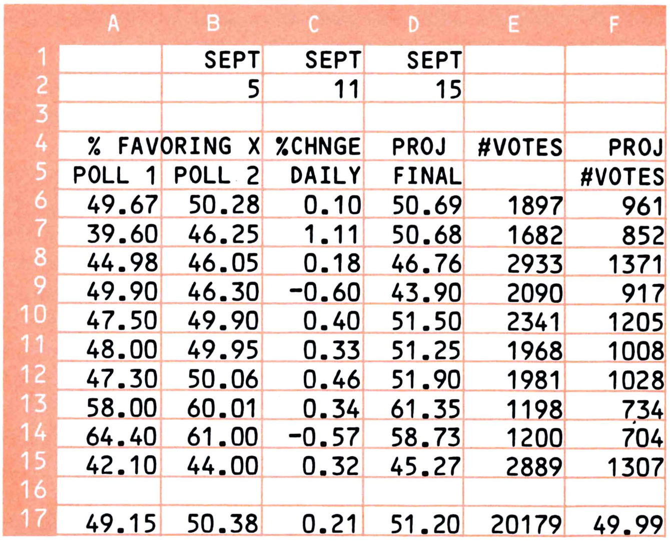
```
>A1:/SLPRCNTS
```

The column showing the number of votes in each precinct should appear in column A, and since each cell in the column was originally formatted with /FI, the values just loaded on the screen should appear as integers, and not in the decimal format.

Since we're concerned with X's projected percentage in each precinct we should now move column A to the right. Enter:

```
>A1:/M . . . F1
```

Enter the labels shown in Figure 4 at the head of column F, and calculate the projected votes X will win.



	A	B	C	D	E	F
1		SEPT	SEPT	SEPT		
2		5	11	15		
3						
4	% FAVORING X		%CHNGE	PROJ	#VOTES	PROJ
5	POLL 1	POLL 2	DAILY	FINAL		#VOTES
6	49.67	50.28	0.10	50.69	1897	961
7	39.60	46.25	1.11	50.68	1682	852
8	44.98	46.05	0.18	46.76	2933	1371
9	49.90	46.30	-0.60	43.90	2090	917
10	47.50	49.90	0.40	51.50	2341	1205
11	48.00	49.95	0.33	51.25	1968	1008
12	47.30	50.06	0.46	51.90	1981	1028
13	58.00	60.01	0.34	61.35	1198	734
14	64.40	61.00	-0.57	58.73	1200	704
15	42.10	44.00	0.32	45.27	2889	1307
16						
17	49.15	50.38	0.21	51.20	20179	49.99

FIGURE 4. *The trend analysis sheet after loading and moving the file PRCNTS, and calculating the PROJ #VOTES column.*

Enter:

```
>F6:/FI@INT(D6/100*E6)
>F6:/R:F7 . . . F15:RR
```

Given the votes in column F, the question is: What percentage of the votes will X win if the trend established between September 5th and September 11th continues? To answer that question add the projected votes in column F and divide by the total in column E. Enter:

```
>F17:@SUM(F6 . . . F15)/
   @SUM(E6 . . . E15)*100
```

That calculation shows X may only win 49.99% of the vote, a smaller margin than the popularity polls alone indicate. And although the aggregate total of the precincts amounts to more than 20,000 voters, the margin of error in the polling results has to be applied to the individual precincts. Given those smaller units, the samples couldn't be large enough to reduce sampling error to 2% or 3%, but is more likely to fall in the range of 10% to 14%. The election may not be as close as the trend of the popularity polls suggests.

One interesting exercise at this stage is to return to row 2 and watch the effects of changing the dates. Making election day one day earlier reduces X's margin; one day later increases it.

As we have seen in both the group weight and trend analysis models, the VisiCalc program allows you to change assumptions easily and watch how those changes in the conditions of the model change the results. Playing "what if" with the winner-takes-all model can be equally informative. And as we showed when we inserted the small part of the group weight model into the trend analysis model, the different models can be combined to provide still more predictive power.

— Paul Hoffman is a freelance writer from California.

VISICALC POINTER . . .

How can you find out the exact value of the entry in a cell?

In the article *Interesting Comparisons* we show that the actual value in one cell is –.000000055. Since the cell was formatted with /F$, the display read 0.00, and the status panel showed the expression beneath the display. There are two ways to find what the actual value is in such instances. The first is to place the cursor on the cell in question and enter /GC and some number of characters that you think might display the entire value. The disadvantages of that approach is that if you are interested in other values nearby, the change in column width may make those cells disappear from the screen, forcing you to scroll around until you find them. The simpler way is to use #. That is, without changing the column width, place the cursor on the cell that you want to examine and enter the symbol #. The status panel will continue to show the expression, the sheet will continue to display the formatted value, but the actual number will appear on the edit line.

To back out of the command you have two choices, each of which produces a very different result and one of which is irrevocable.

If you wish to freeze the value, substituting the number calculated for the expression used to calculate it, press ®.

If you wish to retain the expression in the cell you are examining, use the break key. (On the Apple that is Control C; STOP on the Commodore; and BREAK on the Atari, the Hewlett Packard, the IBM Personal Computer, and the TRS-80 Models I, II, and III.)

Be careful when using #, since pressing ® can wipe out work that you may have spent a lot of time developing.

INTERESTING COMPARISONS

When you are trying to calculate how much you can earn on your money, comparing interest rates compounded at differing frequencies is a commonly encountered problem. One way to compare rates is to search out a table or tables of compound interest and pore over them until you've found the answer. If you can't lay hands on such a table, and you have nice, even numbers and periods to work with, some comparisons can be eyeballed with reasonable accuracy. Using the VisiCalc® program, you can compare rates quickly, cleanly, and with a precision rarely demanded by even a belt and suspenders accountant.

Moreover, even if you only know how much you will start with, how much you will finish with, and how long the money will be out, you can use the technique illustrated in this article to calculate the exact rate you will earn. For example, suppose an investment is proposed in these terms: In return for $10000 today, 167 days hence you will be paid $10867. Assuming that the payment is sound, what rate of interest are you earning on the investment if it is compounded annually? As it happens, you'd earn 19.63%. Daily? 17.93%.

There are three essential questions to answer. First, given a rate that is compounded annually, what are its equivalent rates for daily, monthly, quarterly, and semiannual compounding? Second, given an amount of principal in dollars, an amount that principal will yield, and a timespan, what is the annual rate earned if compounding is calculated daily, monthly, quarterly, and semiannually? And third, without having to search out a table to check our calculations, can we prove that they are correct?

Calculating Rates from a Given Rate

All of the standard financial formulas that deal with compounding, whether present or future values, have one element in common: their compounding periods are expressed as exponents. The opposite of an exponent is a root, and the technique that we will show is based on taking roots.

To begin, set up your sheet with the command:

```
/GC7
```

That creates a seven-character column.

Enter the labels in Column A and across Row 10 as shown in Figure 5.

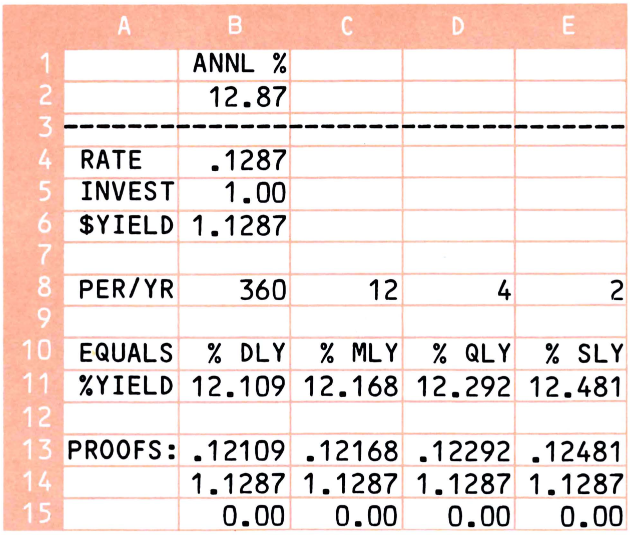
The rate to which we want to find equivalents is entered at B2 beneath the label ANNL %. We entered 12.87 because it is an unusual rate, not often found in many tables. Enter:

```
>B2:/F$12.87
```

At B4 ANNL % is stated as a decimal rate with the entry:

```
>B4:/FG+B2/100
```

Before continuing, some observations about the behavior of money are in order. First, since every dollar



	A	B	C	D	E
1		ANNL %			
2		12.87			
3	-------	-------	-------	-------	-------
4	RATE	.1287			
5	INVEST	1.00			
6	$YIELD	1.1287			
7					
8	PER/YR	360	12	4	2
9					
10	EQUALS	% DLY	% MLY	% QLY	% SLY
11	%YIELD	12.109	12.168	12.292	12.481
12					
13	PROOFS:	.12109	.12168	.12292	.12481
14		1.1287	1.1287	1.1287	1.1287
15		0.00	0.00	0.00	0.00

FIGURE 5. *Although in this figure and the text we have used 360 as the number of days in a year, we could as easily have used 365. The difference between the two amounts to .0000268%, or 27 cents on $1,000,000.*

in any investment compounds at the same rate as every other, the only investment amount that is necessary for us to consider is $1.00. For example, suppose the invested amount is $1234.56. At 12.87% that yields $1393.45, but it still yields only $1.1287 on each dollar. So the only value that is needed to find the equivalent rate is 1.1287. Alternatively, if $1234.56 is invested and yields $1341.98, dividing the yield by the investment produces 1.0870, which is the amount that each dollar invested will produce.

Second, regardless of the number of times money is compounded between the time an investment is made and its yield measured, the rate does not change. That is not to say rates may not change from time to time, as in common stock investments, only that the rate that really matters is the one that produces a difference between the investment and its yield. For example, if you put $10,000 into a common stock, it might grow to $11,000 at the end of six months, and to $12,000 at the end of 18 months. After the first six months its annual rate of growth was 21%, but after 18 months it was 12.92% compounded, in both instances, annually. If you sell at $12,000, it hardly matters that your money once grew at 21%. You'll only realize 12.92% on that investment.

Consequently, because the yield can always be expressed in terms of the yield on $1.00, and the earned rate expressed as a constant over time, the only essential figure that has to be dealt with is the amount to which that dollar will grow at the end of the compounding period.

Although we will not use it in any of the later calculations shown, we have included the amount invested on our sheet in the field INVEST. INVEST is compounded at B6 to give $YIELD. Enter:

```
>B6:+B5*(1+B4)
```

On Row 8 we have entered the number of periods per year for daily, monthly, quarterly, and semiannual compounding. Enter them on your sheet as shown in Figure 5.

The general formula for compounding a single sum is:

$$Y = P * ((1 + i)^n)$$

where Y is the money at the end of the term, P the principal, i the interest rate, and n the number of compounding periods. (Because the number of compounding periods we needed was one, the expression at B6 is the same as if it had been written +B5*((1+B4)^1). You may want to enter that expression somewhere on your sheet to satisfy yourself that you get the same results with both expressions.)

To solve for i, the rate, the general formula is:

$$i = Y / P^{(1/n)} - 1$$

Since the principal, P, equals one, that general formula can be simplified to:

$$i = Y^{(1/n)} - 1$$

If we use 360 for n in that formula, we get the rate by which the principal grows each day (.0003363523). What we want, however, is that amount multiplied by the number of compounding periods in a year, and further multiplied by 100 to give us the annual rate compounded daily and stated as a percent. So enter:

```
>B11:+B6^(1/B8)-1*B8*100
```

That entry is replicated into C11 . . . E11 to produce the equivalent rates for monthly, quarterly, and semiannual frequencies. Enter:

```
>B11:/R:C11 . . . E11:NRR
```

Calculating the Proofs

The simplest way to prove that your calculations are valid is to compound the initial sum by the different periodic rates just calculated. Then, if the rates calculated are correct, the new calculation will produce a value equal to the one calculated at B6.

Since we stated the equivalent rates on row 11 as percents, we must first convert them back into decimals. Enter:

```
>B13:+B11/100®/R:C13 . . . E13:R
```

Now use the decimal rates to compound the investment amount at B5 by the varying rates. Enter:

```
>B14:1+(B13/B8)^B8®/R:C14 . . . E14:RRR
```

Note the interest rate used is the annual rate divided by the number of periods. When that calculation has been made, the result is raised to a power equal to the number of periods, and yields the amount to which the principal will grow at the end of one year. Thus if the values on row 14 are subtracted from B6, we should get zero. Enter:

```
>B15:/F$+B14-B6®/R:C14 . . . E14:RN
```

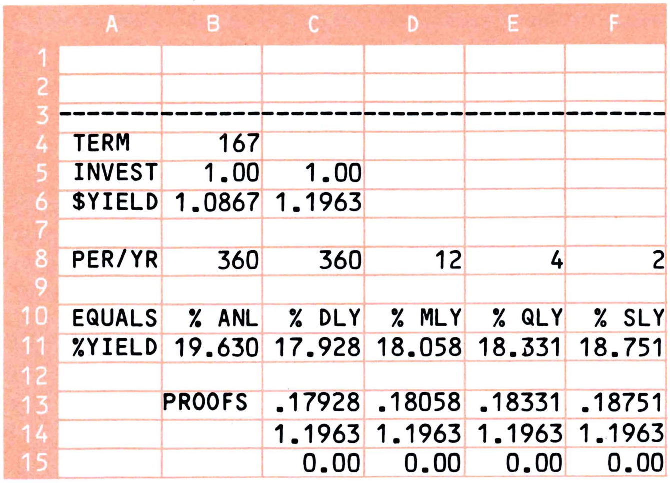
That produces a row of zeroes, which tells us that if the principal is compounded at a different rate and for a different number of periods, it will still amount to the same yield as at the original rate and period. If you change the format from /F$ to /FD, for the default, you will see that there is a very slight difference. Given the rates we have shown, the actual value at B15, for example, is −6.E−8, or −.000000055. The reason for that difference is partly explained by the way that the VisiCalc program handles exponents (by using the log of natural numbers), and partly by the fact that the original rate is defined to only four decimal places, while all subsequent calculations are made to seven or more places. In any event, the difference is not significant.

At this stage, save your sheet for any future calculations that you may wish to make. We will return to it to modify it slightly for solving a slightly different problem, too, so give it the file name CALCALL.

Calculating a Rate from an Unstated Rate

In some transactions we know how much principal is involved, how much the principal will grow to, and the time given for the growth, but not an interest rate. Similarly, the problem may not be a financial one, but may involve, say, population studies in which we measure the size of the population at two different times and want to know the rate at which it has grown.

You should still have the sheet we have developed on your screen, but if not, load the file CALCALL.



	A	B	C	D	E	F
1						
2						
3	----------	----------	----------	----------	----------	----------
4	TERM	167				
5	INVEST	1.00	1.00			
6	$YIELD	1.0867	1.1963			
7						
8	PER/YR	360	360	12	4	2
9						
10	EQUALS	% ANL	% DLY	% MLY	% QLY	% SLY
11	%YIELD	19.630	17.928	18.058	18.331	18.751
12						
13		PROOFS	.17928	.18058	.18331	.18751
14			1.1963	1.1963	1.1963	1.1963
15			0.00	0.00	0.00	0.00

FIGURE 6. *Try entering 1 at B4, .5 at B6, and 1 at B8. That shows the various rates at which you are losing money. The first value calculated, at B11, is the annual loss compounded annually.*

Blank cells B1, B2, and B4, and change the label RATE to TERM as shown in Figure 6. Now insert a column at B1. Enter:

```
>B1:/IC
```

When you blanked B2, all of the cells containing calculations should have been recalculated and should now display either a one or a zero.

In the example cited at the beginning of this article we used $10000, $10867, and 167 days as the principal, yield, and time. If we divide the principal and yield by 10000 we get $1.00 and $1.0867. As shown in Figure 6, enter those values on your sheet in cells B4, B5, and B6.

Since the time over which the investment grows is stated in days, the annual rate compounded annually is the first one to seek. Because we have defined "annual" as 360 days, enter:

```
>B8:360
```

At B11, beneath the label % ANL, we derive the annual rate with an expression similar to the one used above to derive the daily interest. Enter:

```
>B11:+B6^(B8/B4)-1*100
```

If you wish to satisfy yourself that that expression will yield the correct rate, change the values at B4, B6, and B8 to 20, 6.727499, and 1, respectively. The value returned at B11 will be 10%, which can be verified from a compound interest table. (Be sure to change those values back to 167, 1.0867, and 360 before proceeding.)

Now, to convert the annual rate compounded annually to its periodic equivalents, we must calculate the yield on $1.00 at C6 and then proceed exactly as we did above. Enter:

```
>C6:1+(B11/100)
```

That yields the amount to which $1.00 would compound in one year at the rate returned at B11, which should be 19.630%.

The remainder of the sheet will be calculated automatically because the expressions loaded in the file CALCALL are identical to those now needed to convert the annual rate to its equivalents.

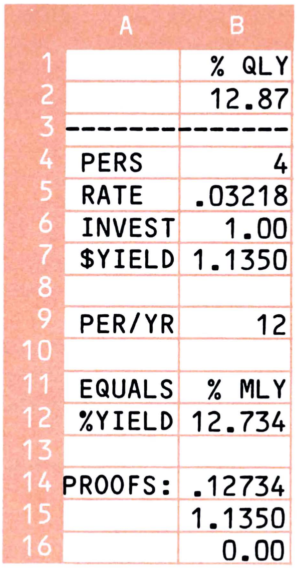
Converting Quarters to Months

Finally, let's consider how to adapt the sheet to the problem of finding, say, the monthly equivalent of a quarterly rate. Save the sheet you have just developed, clear the screen and load CALCALL again. As Figure 7 shows, we have changed the labels slightly. The figure also shows that we have deleted columns C . . . E, and added the field PERS, for periods, on Row 4. You do not have to delete those columns unless you want to run the program a little faster. In any case, you have to add a row to the file, so enter:

```
>B4:/IR4
```

That inserts the new row and enters the number of quarters per year at B4. Since the rate paid will not be 12.87% per quarter, we must adjust the expression at B5 accordingly. You can use /E to change the expression or enter the appropriate expression from scratch. In either case, the expression at B5 should read:

```
>B5:+B2/100/B4
```



	A	B
1		% QLY
2		12.87
3	----------	----------
4	PERS	4
5	RATE	.03218
6	INVEST	1.00
7	$YIELD	1.1350
8		
9	PER/YR	12
10		
11	EQUALS	% MLY
12	%YIELD	12.734
13		
14	PROOFS:	.12734
15		1.1350
16		0.00

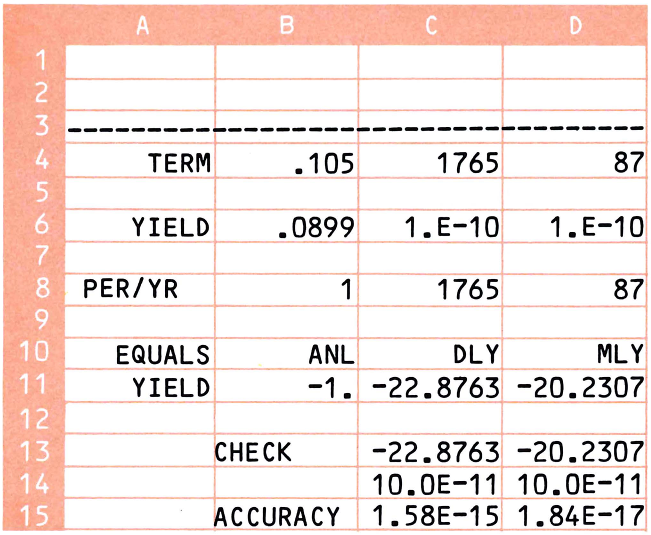
FIGURE 7. *By changing the entry at B8 from 4 to 1 you convert a rate compounded quarterly into its equivalent rate compounded annually.*

INVEST remains the same, but we must edit $YIELD to reflect the changes made above. Once again, use /E to make the necessary change, or enter:

```
>B7:1+B5^B4
```

Since 360 is still on the sheet from CALCALL, the rate that now appears at B12, 12.670, is the daily rate that is equivalent to a 12.87% quarterly rate. By changing the value at B9, you are able to calculate any equivalent rate. Change it to 12 and B12 will return 12.734, while entering 1 at B9 will produce the annually compounded rate, 13.505%. And, of course, if B9 is changed to 4, B12 will return 12.87%, the rate shown at B2.

Although we have concentrated on financial comparisons, the techniques shown in this article can be applied to comparisons of any growing population, whether it is a population of dollars, people, or bacteria in a petri dish. Within limits, it can also measure comparative rates of shrinkage. That is, if the $YIELD, or its analogous entry, is less than $1.00, the expressions we have shown will give the rate at which the population is declining. However, if you couple a very large decrease with a very short period of time, the expressions we have used will not be able to calculate the rate at which the population is declining. If the application that you have in mind for the points made in this article must calculate very large declines (say from 1 to .1 in 1 period), use the expression shown in the caption to Figure 8.



	A	B	C	D
1				
2				
3	----------	----------	----------	----------
4	TERM	.105	1765	87
5				
6	YIELD	.0899	1.E-10	1.E-10
7				
8	PER/YR	1	1765	87
9				
10	EQUALS	ANL	DLY	MLY
11	YIELD	-1.	-22.8763	-20.2307
12				
13		CHECK	-22.8763	-20.2307
14			10.0E-11	10.0E-11
15		ACCURACY	1.58E-15	1.84E-17

FIGURE 8. *This sheet is an edited version of the one shown in Figure 6. The expressions used, however, are very different, and allow manipulation of extreme cases in which the rate of decline is very high and the values involved very small. The expression at C6 is* 1+B11, *not* 1+(B11/100) *as in the text. On this sheet, none of the expressions that are carried from the sheet developed in the text are multiplied by 100 to generate percentage values. At B11, enter* >B11:@EXP(@LN(B6)/B4)−1*B8. *Note we've added values in C4, D4, C8, and D8. Enter them, then replicate the expression at B11 into C11 . . . D11. The redundancies on rows 4 and 8 allow us to replicate the expression at B11 making all value references relative.*

SUBMITTING ARTICLES TO SATN

Articles and pointers submitted to SATN are considered for acceptance based on their quality and interest to SATN readers.

We would prefer to see queries first. If you submit an article without querying us first, include a stamped, self-addressed envelope (or mailer for diskettes). If you do not enclose an SSAE or mailer, we cannot return your material. Prospective authors may send for the *SATN Writer's Guide.*

Address all queries, articles, or requests for the guide to:

SATN
Editorial Department
P.O. Box 494
Cambridge, MA 02139

TEE SHIRTS, DRIVE SHAFTS, AND BONDS

Consider the problem faced by a company marketing a small line of tee shirts. Its entire product line consists of only two printed designs, one featuring a puppy, the other a kitten. But the shirts are available in two styles, adult's and children's; five sizes, x-small, small, medium, large, and x-large; five colors of shirt; and three colors of ink. When all the different possible combinations of design, style, sizes, and colors are taken into account, how many products does the company really have? Two? Ten? A few hundred?

Well, if any tee shirt can come in any size, color, style, and so forth, the company's product line effectively consists of 300 different products. How did we arrive at that number? By multiplying the number of elements in each category by the number of elements in each of the other categories.

*(2 Designs * 2 Styles * 5 Sizes * 5 Colors * 3 Ink Colors)*

There are no such things as children's extra large or adult's extra small tee shirts, so the company's product line is smaller, only 240 products (2*2*4*5*3). But that's not much smaller, so the key problem remains, and that problem is keeping track: of orders received, of orders to suppliers, printing orders, puppy shirts, kitten shirts, adult's, children's, large, medium, ad infinitum.

The first step toward keeping track is to assign every different combination of elements its own, unique number; a descriptive code. Once that is done it becomes possible to exercise the kind of control that can make the difference between profit and loss.

There are many ways to create such code numbers. The most obvious, in the example we've used, is to make up a five-digit code, the first digit of which stands for the design, the second for style, and so forth. The method that we shall use requires only three digits, and allows you to increase the number of elements to some degree without increasing the number of digits in the code equally.

As our title suggests, tee shirts are one example of the things to which a descriptive code number might be applied. Stocks and bonds might be coded in terms of the companies that issue them, the year of maturity, coupon rate, whether they are convertible, preferred, and so on. Information about a drive shaft might include the material of which it is made, the manufacturer(s) who make it, the machine or subassembly into which it fits, and the location to which it should be shipped.

The technique we shall describe can also be used to analyze market research or polling data, or in any application in which it is useful to consolidate disparate bits of information into a single easily manipulated number. Once the identifying code number is created, any individual element can be retrieved.

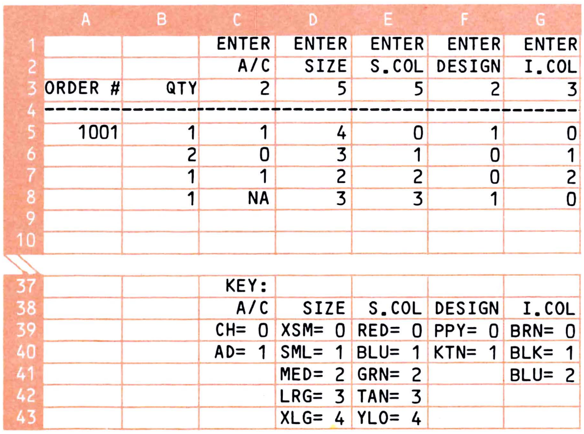
Creating a Code Number

Begin by setting up your sheet with the commands to calculate by rows and set the column width to seven characters. Enter:

```
/GOR
/GC7
```

Because we are dealing with theoretical orders for tee shirts we have entered the labels ORDER # and QTY in A4 and B4. Enter those labels and the others shown in Figure 9 in C2 . . . G2 and C3 . . . G3. Beneath the labels, in C4 . . . G4, enter the values shown in the figure. Those values are the maximum number of variations in each category. For example, A/C refers to adult or children's shirts, so there are only two variables.

Coding begins on row six. In any category the value assigned to a characteristic may range from zero to one less than the maximum entered on row 4. Thus, the values that may be entered in column C may be either zeroes or ones. In column D, however, the range may be 0 . . . 4. To enter the values that will be coded, or coding values as we'll call them, you can either use a printed paper key showing the variables in each column, or construct a key elsewhere on the VisiCalc® sheet, as we have done (see Figure 9).



	A	B	C	D	E	F	G
1			ENTER	ENTER	ENTER	ENTER	ENTER
2			A/C	SIZE	S.COL	DESIGN	I.COL
3	ORDER #	QTY	2	5	5	2	3
4	-------	-------	-------	-------	-------	-------	-------
5	1001	1	1	4	0	1	0
6		2	0	3	1	0	1
7		1	1	2	2	0	2
8		1	NA	3	3	1	0
9							
10							
37			KEY:				
38			A/C	SIZE	S.COL	DESIGN	I.COL
39			CH= 0	XSM= 0	RED= 0	PPY= 0	BRN= 0
40			AD= 1	SML= 1	BLU= 1	KTN= 1	BLK= 1
41				MED= 2	GRN= 2		BLU= 2
42				LRG= 3	TAN= 3		
43				XLG= 4	YLO= 4		

FIGURE 9. *After splitting the screen (>A11:/WH) and moving the cursor to the lower window with the semicolon, you can enter an abbreviated key to the entries that must be made above. The key shown is made up entirely of labels.*

After entering the coding values for the first three tee shirts as shown in Figures 9 and 10, enter @NA in C9, and the values shown in the figure in D9 . . . G9. That row says that one of the tee shirts on order 1001 is not identified as either an adult's or child's shirt. Rather than not posting the order, we have chosen to post it and flag the missing information with @NA. We cannot simply leave the cell blank, because the method used to create the code reads blanks as zeroes. Since zero is a valid entry, a blank cell will yield an incorrect code number. Therefore, if one element of information is missing, we must either enter none at all on that line or use @NA.

To the right of the area of the sheet where the coding values are entered, the actual coding itself takes place. Begin the CODE section with the entry:

```
>C3:/R: . . . G3:I3:
```

That brings the labels to the right. Then enter:

```
>C4:/R: . . . G4:I2:
```

That brings the maximum values for each category into row 2 of the CODE section. It is not absolutely necessary to carry the maximum values into this section of the sheet, but should you want to create separate files for entering data, coding it, and breaking down the codes, they are much easier to construct if the data is repeated on each part of the sheet we are developing.

We showed earlier that there are 300 possible combinations of style, size, color, ink, and design. That number is the product of multiplying the maximums in each category. To code each tee shirt separately, we will generate a series of intermediate products, which we shall call "running products," and use them to multiply the coding values in the corresponding columns. We begin by entering a 1 in I4 and multiplying it by the maximum shown in I2. Enter:

```
>I4:1
>J4:+I4*I2
```

The remaining running products are produced by replicating the formula in J4 into K4 . . . N4. Enter:

```
>J4:/R:K4 . . . N4:RR
```

If you have entered the values shown in Figure 9, the preceding entry should produce the series 1, 2, 10, 50, 100 and 300 in cells I4 . . . N4.

The first coding value is now coded by multiplying it by the running product in row 4 of the corresponding column. Thus, the zero that identifies a child's tee shirt is multiplied by one. Enter:

```
>I6:+I4*C6
```

By replicating that entry into the columns to the right, the rest of the coding values are coded for each category. Enter:

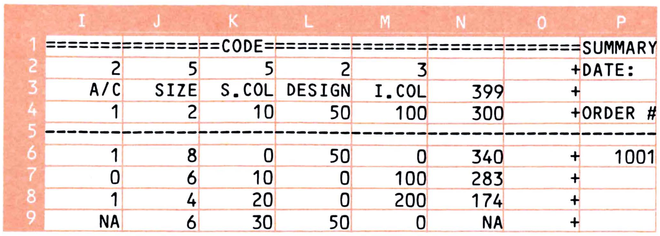
```
>I6:/R:J6 . . . M6:RR
```

To encode the entries in rows 7 . . . 9, replicate I6 . . . M6 into the rows below, keeping references to the values in row 4 as N (No Change) and to row 6 as R (Relative). Enter:

```
>I6:/R . . . M6:I7 . . . I9:NRNRNRNRNR
```

Notice that I9 displays NA. All future references to that cell will also return NA until a value is entered at C9.

Now, to produce the unique number that describes the specific combinations entered in columns C . . . G, we could add the products in each of the rows that were produced by the last entry. However, if five consecutive zeroes are entered in C6 . . . G6, the code number that would appear in N6 would also be zero. In fact, as it



	I	J	K	L	M	N	O	P
1	=========	=========	==CODE=	=========	=========	=========	=========	SUMMARY
2	2	5	5	2	3		+	DATE:
3	A/C	SIZE	S.COL	DESIGN	I.COL	399	+	
4	1	2	10	50	100	300	+	ORDER #
5	---------	---------	---------	---------	---------	---------	---------	---------
6	1	8	0	50	0	340	+	1001
7	0	6	10	0	100	283	+	
8	1	4	20	0	200	174	+	
9	NA	6	30	50	0	NA	+	

FIGURE 10. *Repeating the maximums of each category in row 2, rather than referring to the values in columns C . . . G, row 4, makes it possible to break down the sheet into smaller, separate files for data entry, encoding, summarizing, and decoding.*

now stands the coding mechanism will produce 100 part numbers that are one or two digits long. That is, the series from 0 . . . 99. But we can add an entry to the sheet that will force all codes to be at least three digits long. Enter:

```
>N3:399
```

That is the smallest number that will force all codes to be three digits long. You could in fact use any number between 399 and 999.

Now enter:

```
>N6:+N3-@SUM(I6 . . . M6)
```

Replicate that entry into N7 . . . N9. Enter:

```
>N6:/R:N7 . . . N9:NRR
```

The sums produced by that entry are the code numbers that uniquely identify the particular combination of variables entered in rows 6 . . . 9, columns C . . . G.

Since we are keeping track of tee shirt orders, we have added a SUMMARY section to the sheet, located immediately to the right of the CODING section. The purpose of the SUMMARY is to produce a printable report for order fulfillment and control, and billing. Though the applications to which you would put a coding system to work may not require a summary, we suggest you go through the following section to see how @ISNA is used.

Summarizing

After entering the labels shown in Figure 11, the order number and quantity of each tee shirt ordered are carried into the summary with the entries:

```
>P6:+A6
>Q6:+B6
```

The QTY column is created by replicating the entry at Q6 down the sheet relatively. Enter:

```
>Q6:/R:Q7 . . . Q9:R
```

To generate the ITEM column, replicate the code numbers that appear in column N. Since that column is being copied exactly, all value references are N (No Change). Although we show only three 'N' references in the following expression, you will have to enter three for each line into which the expression is being replicated. Enter:

```
>N6:/R . . . N9:R6:NNN
```

The PRICE column is created by using an @LOOKUP table, which we have located well below the SUMMARY. Since the table does not change, it can be located anywhere on the sheet and not create a problem of forward references. Figure 11 shows a split screen, and the location of the table on rows 38 and 39, columns R and S. Enter the table as shown in the figure.

Notice that the prices shown in column S do not display NA, even though the ITEM column does. The @SUM function of the VisiCalc program will be used in column T to find the total of the column. If we allowed NA to appear in the column being added, it would give us NA as the total. In order to get at least a partial total, we can use @IF and @ISNA to produce a zero or a price which can then be used by @SUM. Enter:

```
>S6:/F$@IF(@ISNA(R6),0,
  @LOOKUP(C6,R38 . . . S38))
>S6:/R:S7 . . . S9:RRNN
```

Calculating EXT., the extension of the order, is a matter of multiplying the quantity of an item by its price. Enter:

```
>T6:/F$+S6*Q6®/R:T7 . . . T9:RR
```

To total an order we use @SUM placed alongside the last entry for the order. Enter:

```
>U9:/F$@SUM(T6 . . . T9)
```

After returning the cursor to P1, you can create a print file with /PF, define the lower right at the last entry in column U, and generate a printed report.

Imagine that 15 or 20 orders have been entered, coded, and summarized, and you now have a summary sheet that is 100 rows deep. Using DIF, you can create a file with /S#S and sort the code numbers and the quantities ordered by transferring the DIF file to a database management program that supports DIF. Once the sort is complete you can return the file to the VisiCalc program, again via DIF, and analyze the various quantities of different sizes, designs, and colors of shirts or other variables.

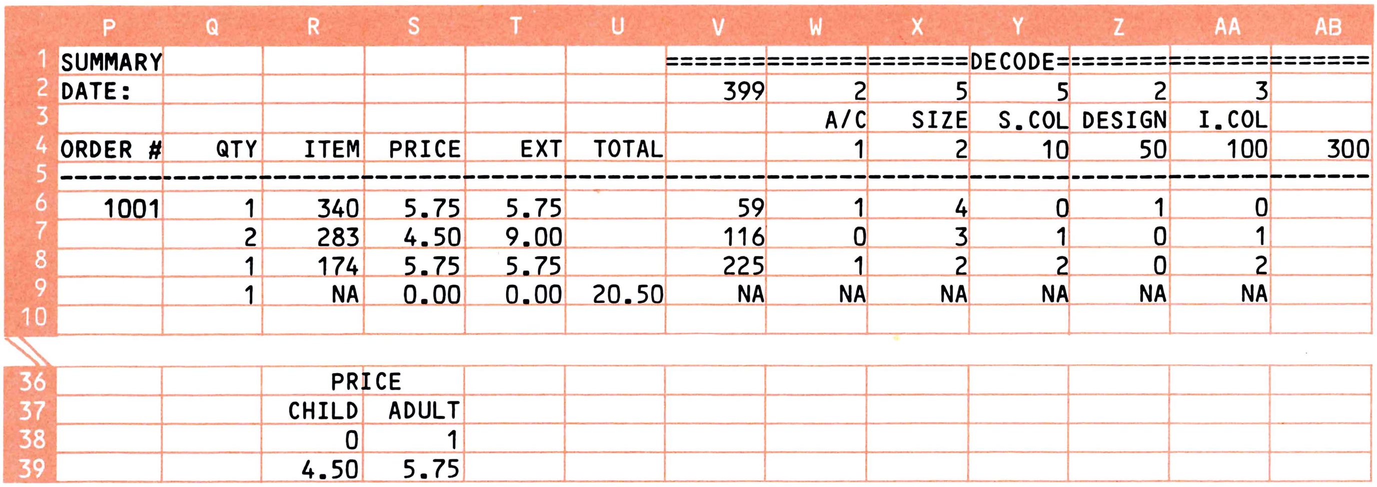
Breaking Down the Identifier

A code is useless unless it can be read by those for whom it's intended, and so the next step is to break down the identifier into its component parts. (That's especially the case if you have performed a sort.) Breaking down the code number is the reverse of building it and requires, on the sheet we have developed, the same values and labels as the coding section. Thus, all that need be done to get started is to replicate rows 2, 3, and 4 from the coding section at the right. Enter:

```
>I2:/R . . . M2:W2:
>I3:/R . . . M3:W3:
>I4:/R . . . N4:W4:RRRRRRRRRR
```

Compare Figures 10 and 11.

The last step performed in creating the code numbers for each tee shirt was the subtraction of the sums of I6 . . . M6 through I9 . . . M9 from 399. Therefore, the



	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	SUMMARY						==========	==========	=========	DECODE====	==========	==========	==========
2	DATE:						399	2	5	5	2	3	
3								A/C	SIZE	S.COL	DESIGN	I.COL	
4	ORDER #	QTY	ITEM	PRICE	EXT	TOTAL		1	2	10	50	100	300
5	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------
6	1001	1	340	5.75	5.75		59	1	4	0	1	0	
7		2	283	4.50	9.00		116	0	3	1	0	1	
8		1	174	5.75	5.75		225	1	2	2	0	2	
9		1	NA	0.00	0.00	20.50	NA	NA	NA	NA	NA	NA	
10													
36			PRICE										
37			CHILD	ADULT									
38			0	1									
39			4.50	5.75									

FIGURE 11. *The Summary and Decoding portion of the sheet. Note that if you transfer data from the sheet to a database management program, it might be wise to repeat the order number for each entry so that after a sort by code number you can still tie each code to the appropriate order.*

first step in breaking down the codes is the subtraction of the codes from that value. Enter:

```
>V2:399
>V6:+V2-R6®/R:V7 . . . V9:NR
```

Now we can retrieve the individual elements from the values at V6 . . . V9. We built the code by multiplying the coding values by the running products. Breaking it down calls for dividing the codes by the running products. The answers that those operations produce are not quite what we are after, though. The value we need is not the number of times that the code number can be divided by a running product, but the remainders of any divisions. Each remainder, when divided by the next running product, yields the coding value. Because we built up the code by applying running products from left to right, we break it down by applying them from right to left.

In more mathematical terms, we must break down the code using modulo, an operation that yields the remainder of division as its answer, rather than a quotient made up of an integer and a decimal portion. As we have seen in earlier issues of SATN (*Carrying on with VisiCalc*, Vol. I, No. 2; and *A General Formula for Date Arithmetic*, Vol. I, No. 3), a modulo function can be performed in the VisiCalc program with @INT. The general formula is:

$$\text{Value} - \text{@INT}(\text{Value} / n) * n) = \text{Remainder}$$

where n is the running product in the column to the right of the one in which the calculation is made. Because we must further divide the results of that operation by the running product applied directly to the original coding value, the full general expression is:

$$\text{@INT}(\text{Value} - (\text{@INT}(\text{Value} / n_1) * n_1)/ n_2)$$

where n_1 is the running product in the column to the right and n_2 the running product in the column being calculated.

The actual expression to break the code is:

```
>W6:@INT(V6-(@INT(V6/X4)*X4)/W4)
```

Replicate that expression down the sheet with the entry:

```
>W6:/R:W7 . . . W9:RRNNN
```

Then replicate that column into the others with the entry:

```
>W6:/R . . . W9:X6 . . . AA6:NNRRR
```

repeating the value references (NNRRR) five times.

Although we have used a single sheet, each major part of the problem can be broken down into a smaller sheet. For example, entering the coding values can be confined to a single sheet that can be saved and loaded onto another sheet that would begin at I1. That second sheet could contain the coding section and summary. A third sheet could be used only to break down the codes, and could easily be relocated from its position as shown in Figure 11 and loaded with a single column containing only code numbers that need breaking down. While applications like the one shown might not require separate sheets, others may require more data security and breaking things down into pieces is one way of providing such security.

BACK ISSUES AVAILABLE

Back issues of SATN are now available. The cost is $6.00 per copy for orders mailed within the USA. If your order is to be mailed outside the USA, add $2.50 for shipping and handling to the full price of the order.

To order back issues, please send your name and address, a list of the specific issues of SATN you want (by Volume and Issue numbers), and a check or money order payable to Software Arts, Inc., to:

SATN Back Issues
P.O. Box 494
Cambridge, MA 02139